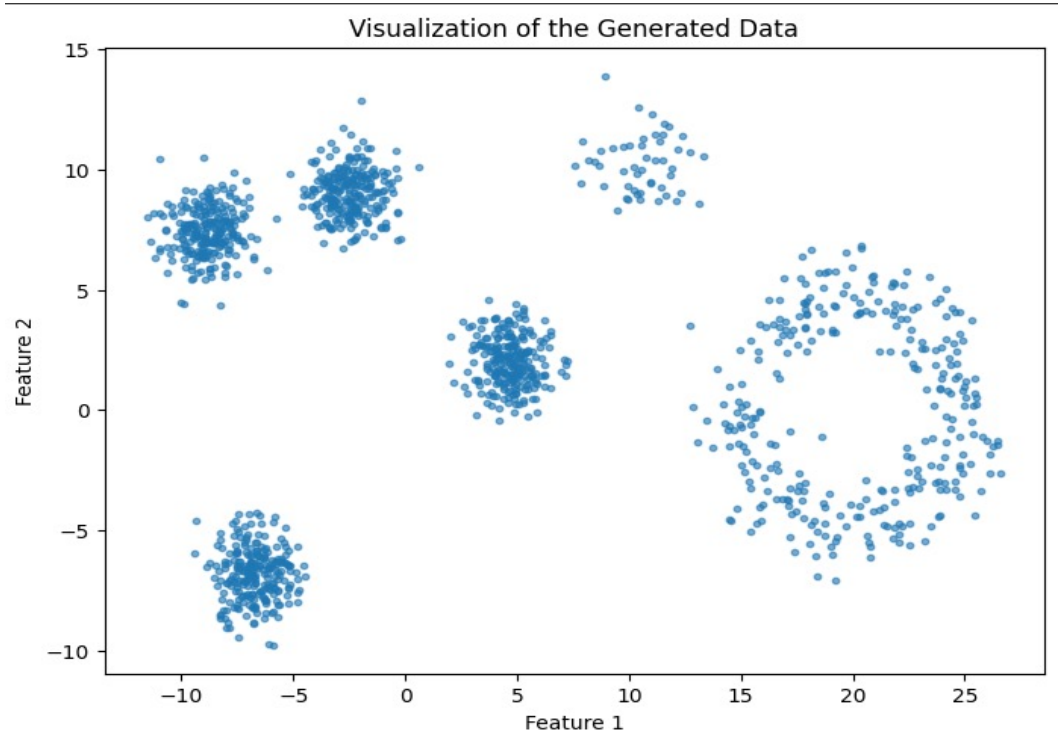
Machine learning – exe 3

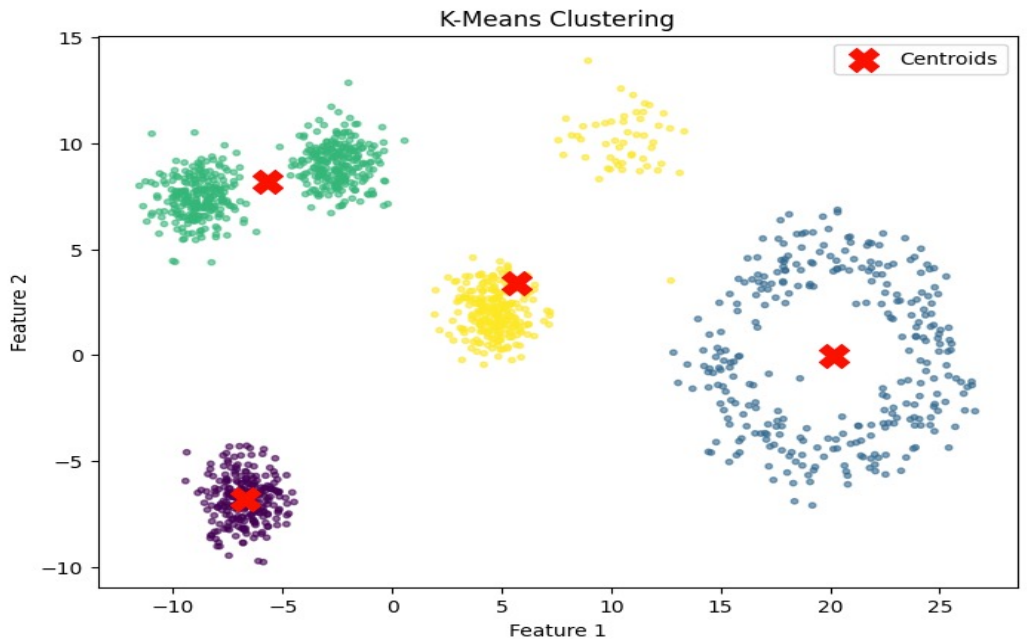
Zaid habiballah - 322513433

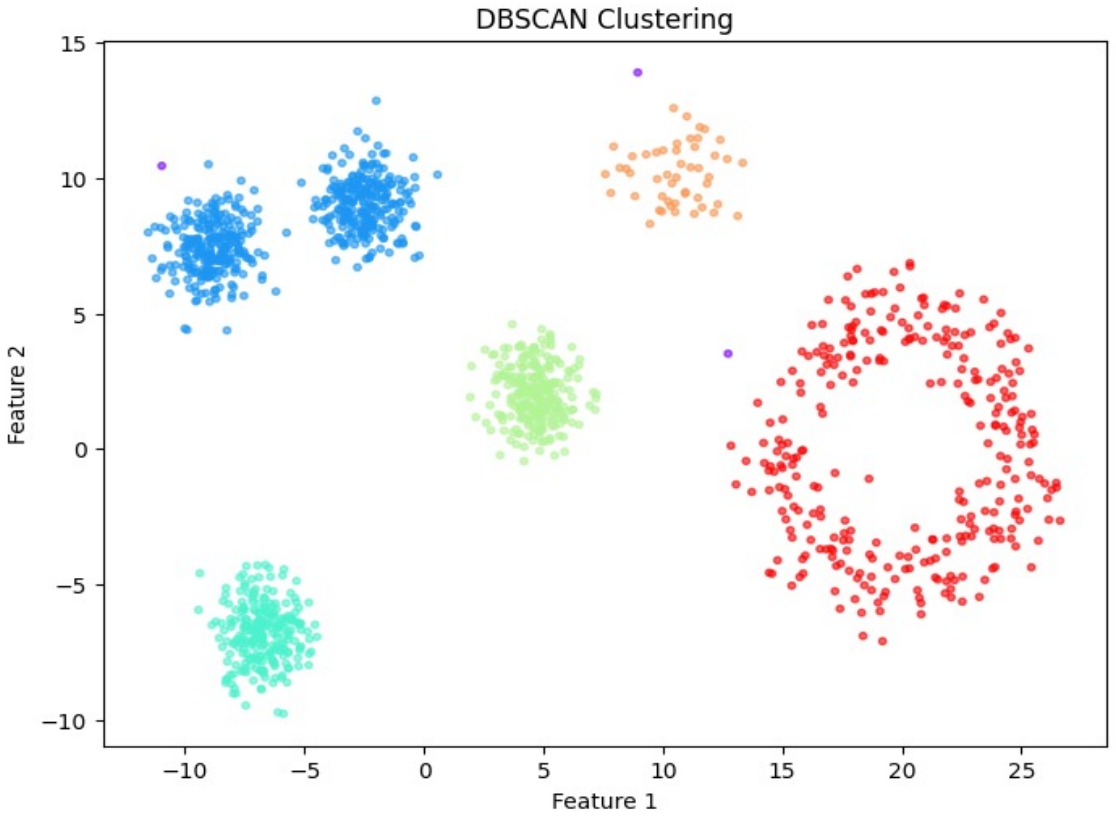
The first data

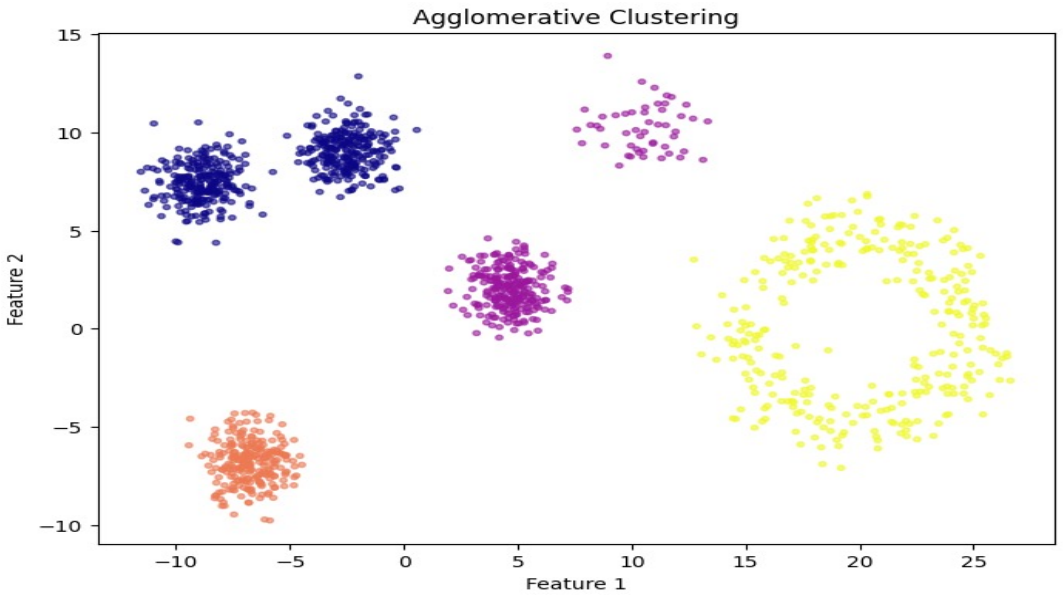
```python
import numpy as np
import matplotlib.pyplot as plt

# Call create_data
X = create_data(1.0, 1000, 50, 300, 200)

# Data visualization
plt.figure(figsize=(8, 6))
plt.scatter(X[:, 0], X[:, 1], s=10, alpha=0.6)
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.title("Visualization of the Generated Data")
plt.show()
```



Try 1 – with 4 clusters and eps 1.5

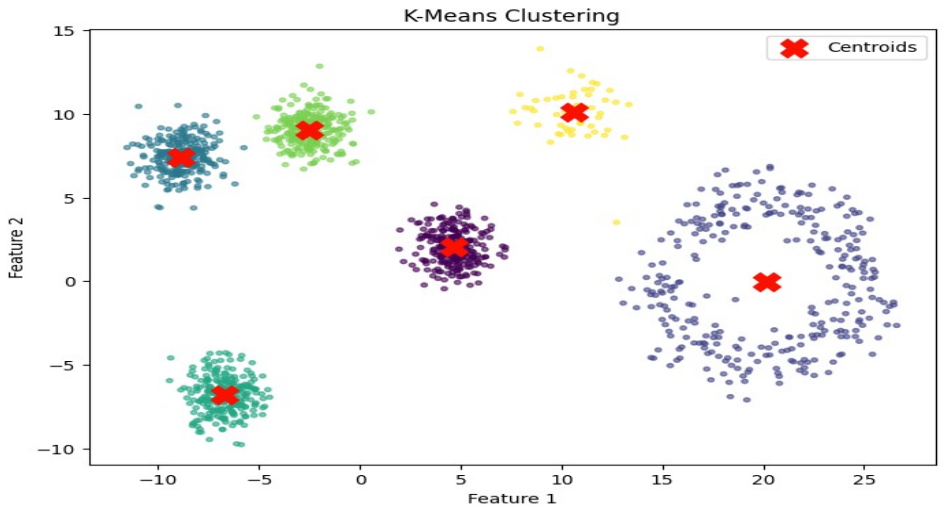


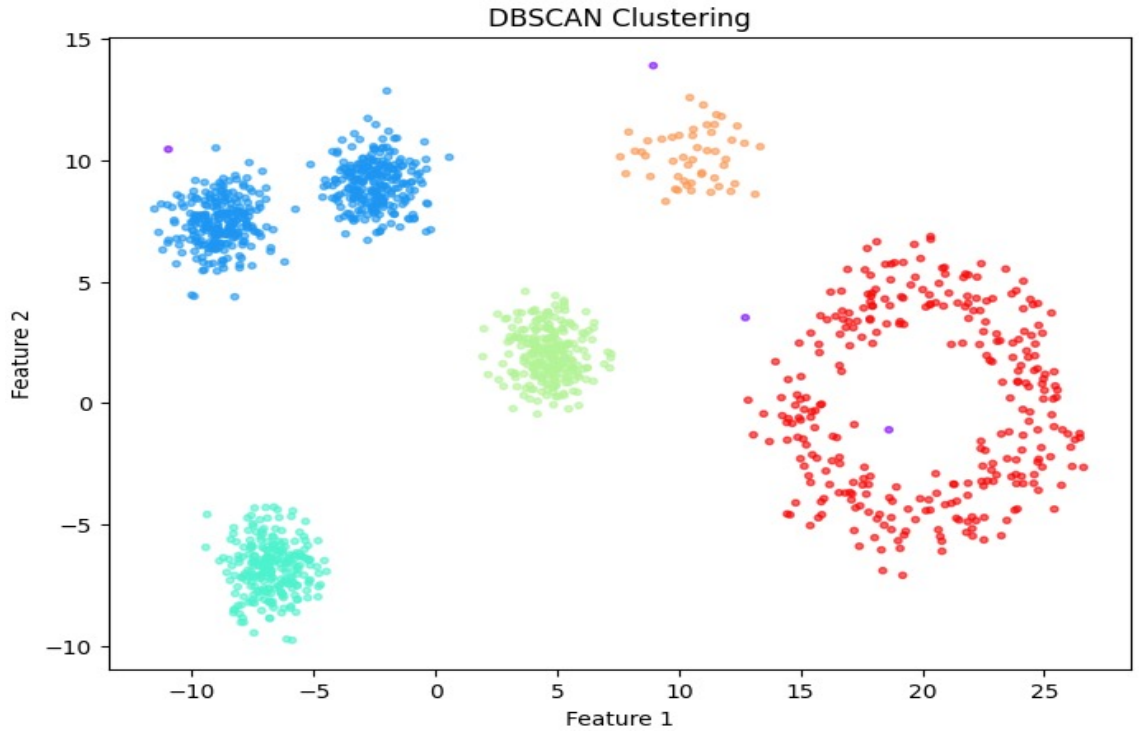


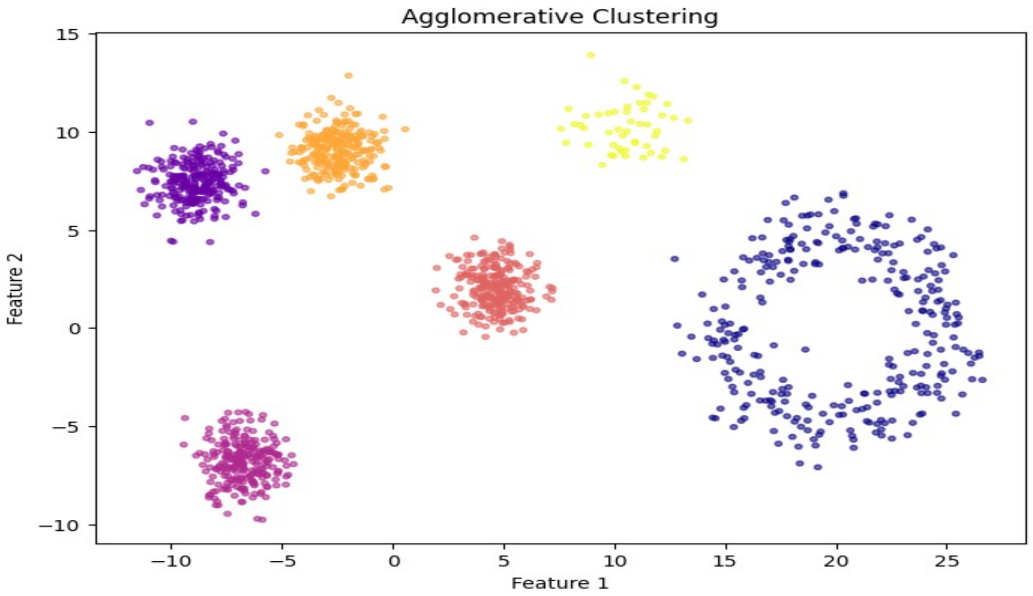



```
Silhouette Score Comparison of Clustering Models:

K-Means: Silhouette Score = 0.6813
DBSCAN: Silhouette Score = 0.6894
Agglomerative: Silhouette Score = 0.6813

Best model based on Silhouette Score: DBSCAN
```

Try 2 – with 6 clusters and eps 1.5

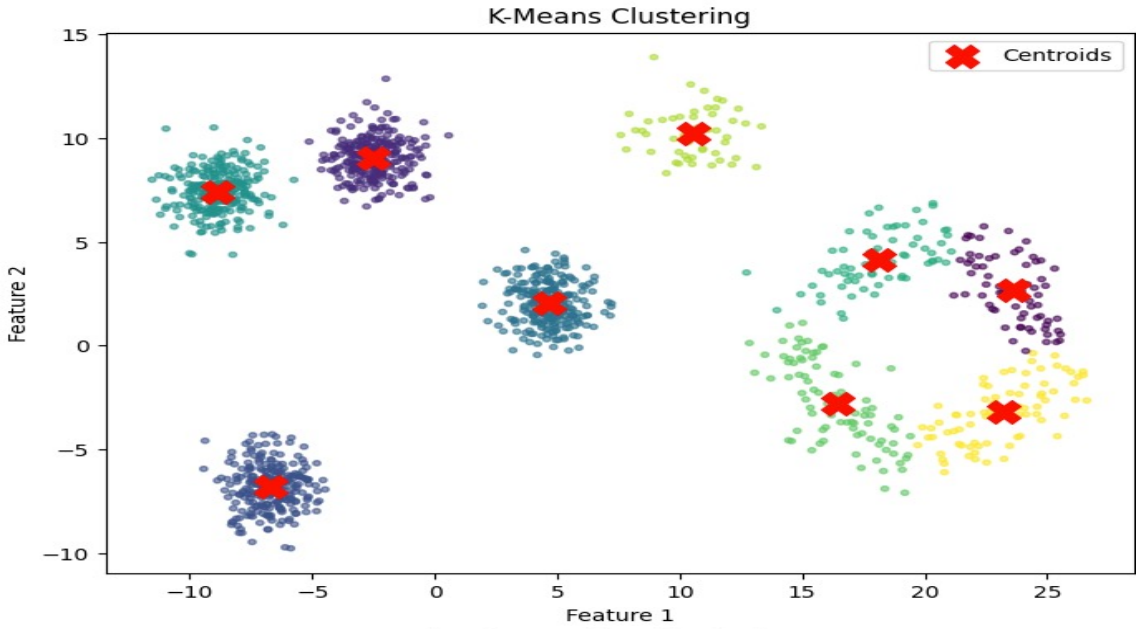


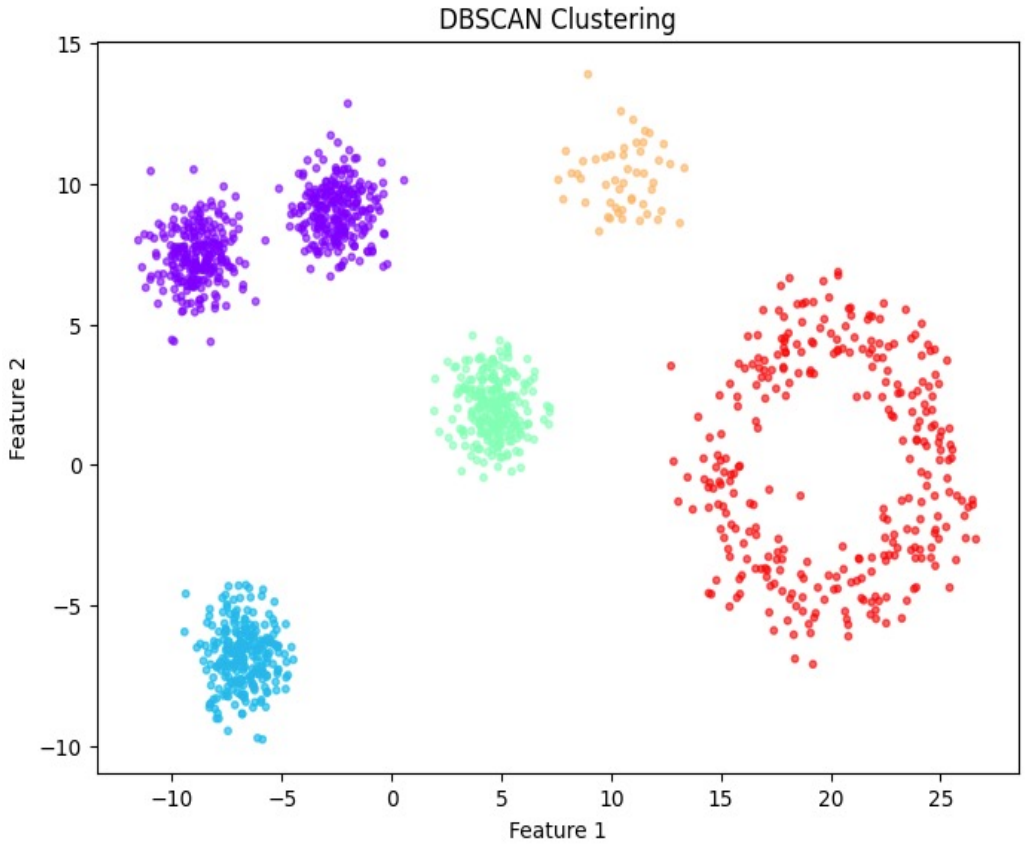


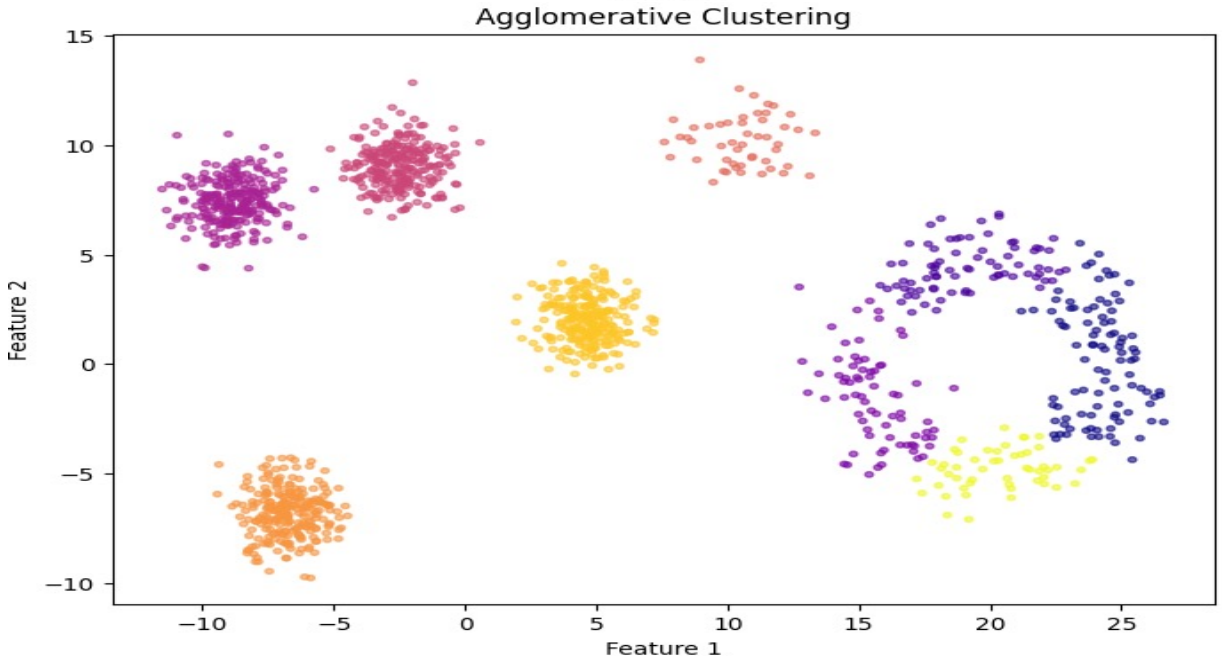



```
Silhouette Score Comparison of Clustering Models:

K-Means: Silhouette Score = 0.7251
DBSCAN: Silhouette Score = 0.6888
Agglomerative: Silhouette Score = 0.7257

Best model based on Silhouette Score: Agglomerative
```

Try 3 – with 9 clusters and eps 2.5



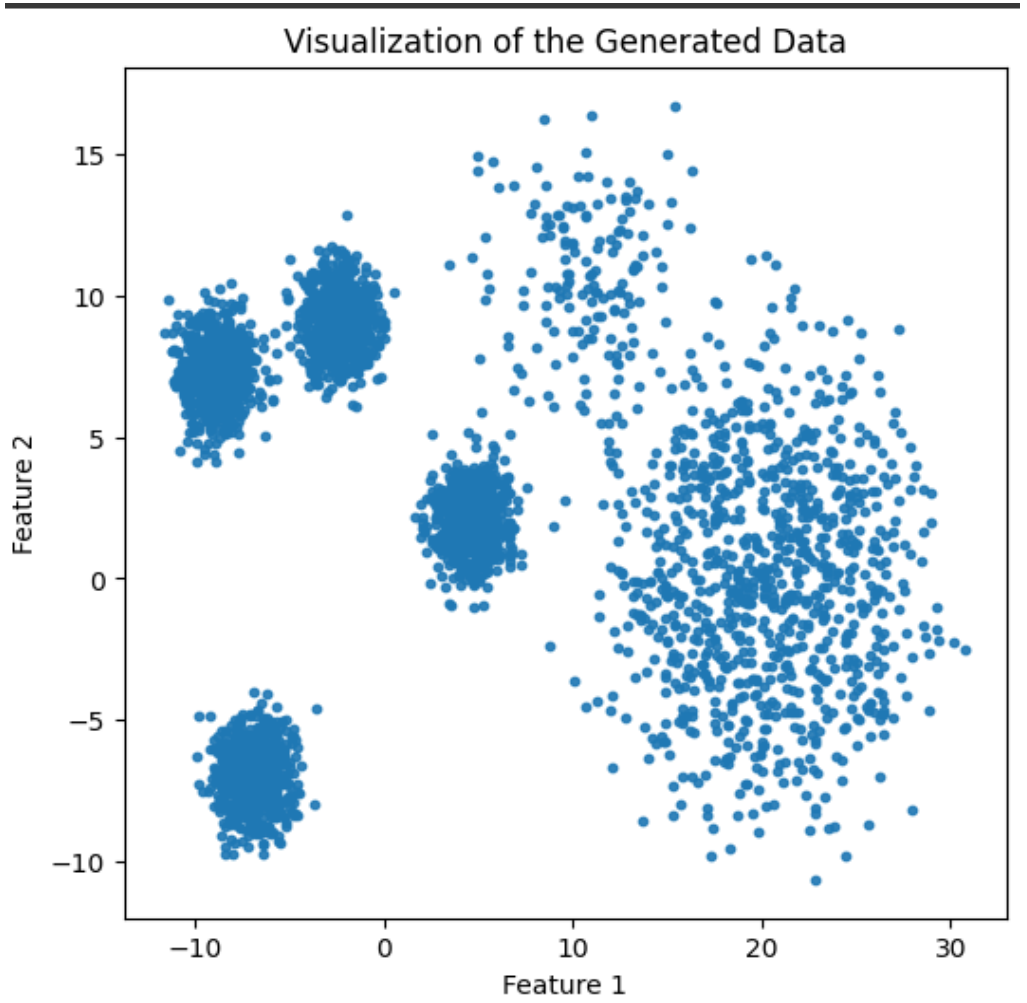




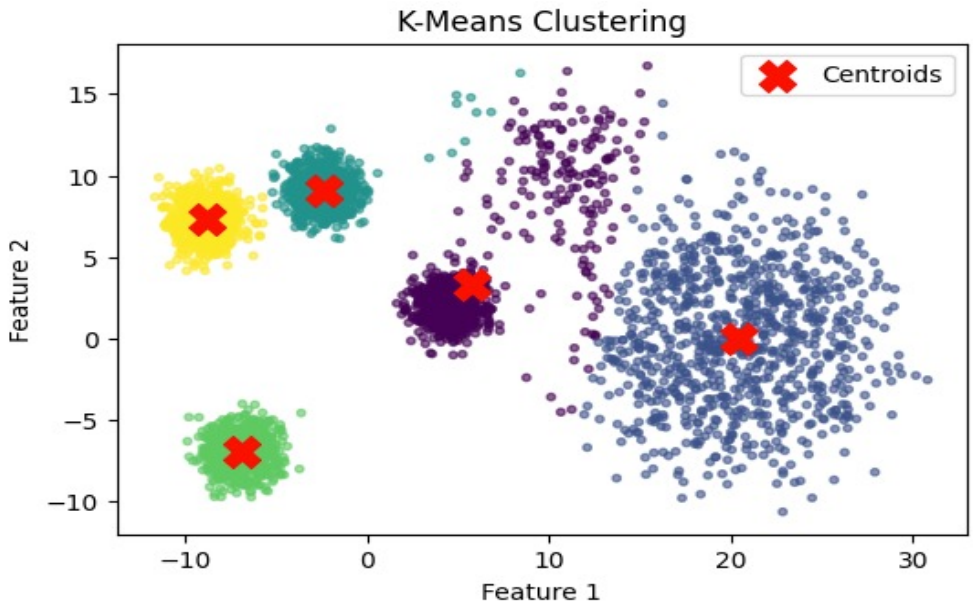

```
K-Means: Silhouette Score = 0.7161
DBSCAN: Silhouette Score = 0.6923
Agglomerative: Silhouette Score = 0.7088

Best model based on Silhouette Score: K-Means
```

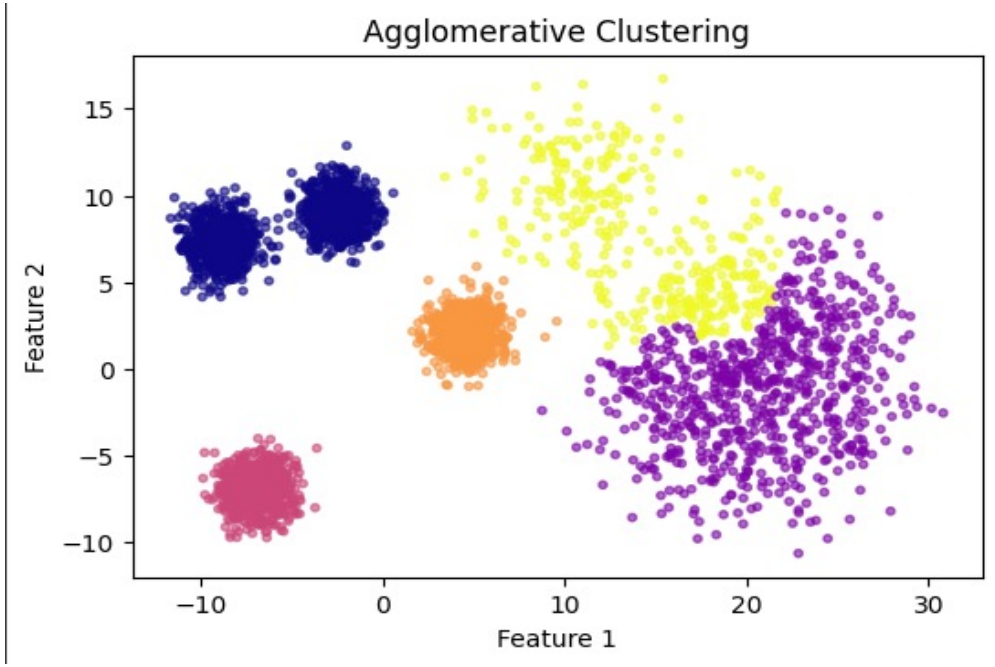
The second data

```python
import numpy as np
import matplotlib.pyplot as plt

# Call create_data
X = create_data(2.5, 3000, 150, 900, 750)

# Data visualization
plt.figure(figsize=(6, 6))
plt.scatter(X[:, 0], X[:, 1], s=10, alpha=0.9)
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.title("Visualization of the Generated Data")
plt.show()
```

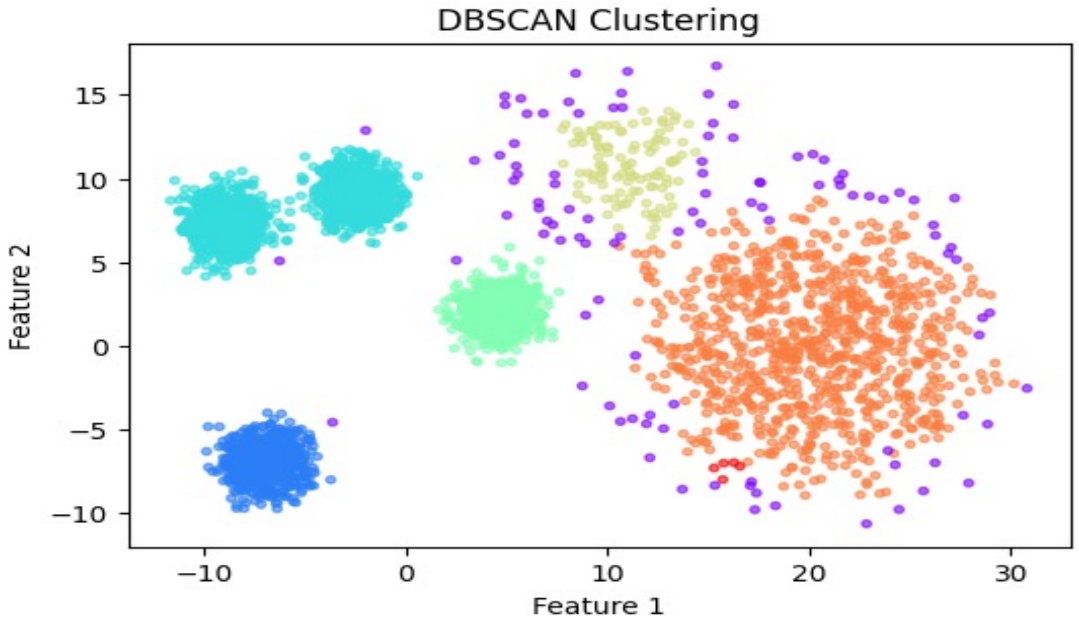


Try 4 – with 5 clusters and eps 1

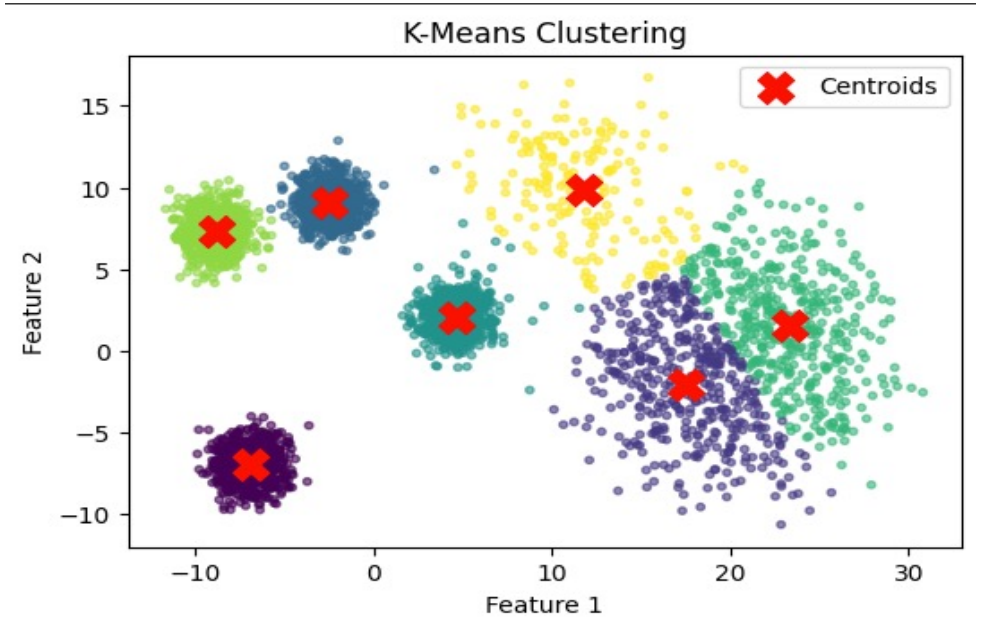


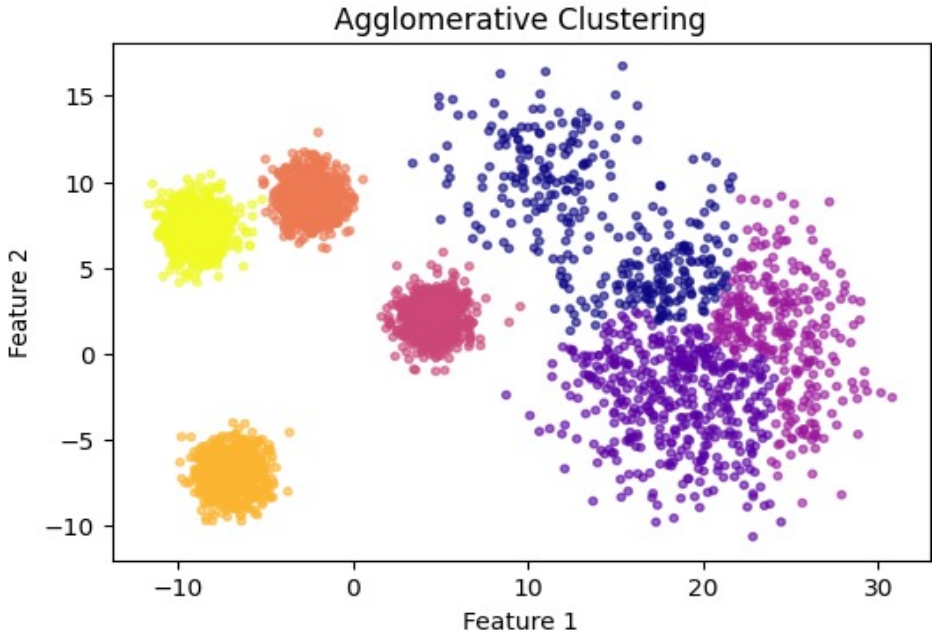


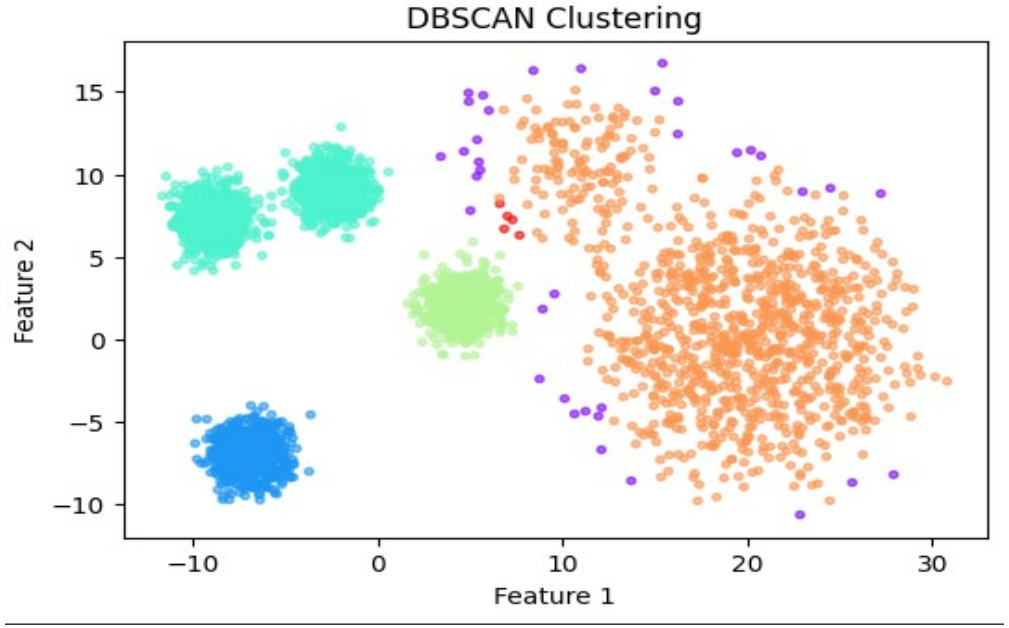

```
Silhouette Score Comparison of Clustering Models:

K-Means: Silhouette Score = 0.6812
DBSCAN: Silhouette Score = 0.5895
Agglomerative: Silhouette Score = 0.6459

Best model based on Silhouette Score: K-Means
```



Try 5 – with 7 clusters and eps 1.5








```
Silhouette Score Comparison of Clustering Models:

K-Means: Silhouette Score = 0.6755
DBSCAN: Silhouette Score = 0.5883
Agglomerative: Silhouette Score = 0.6640

Best model based on Silhouette Score: K-Means
```